AD658068

TECHNICAL PAPER
RAC-TP-262
JULY 1967

Theory of Lagrange Multipliers for Constrained Optimization Problems

by
James E. Falk

DISTRIBUTION STATEMENT
"This document has been approved for public release and sale; its distribution is unlimited."

PREPARED FOR THE
DEPARTMENT OF THE ARMY
Contract No. DA 44-188-ARO-1

Copy 28 of 185

31

DISTRIBUTION STATEMENT
"This document has been approved for public release and sale; its distribution is unlimited."

ADVANCED RESEARCH DEPARTMENT
TECHNICAL PAPER RAC-TP-262
Published July 1967

Theory of Lagrange Multipliers for Constrained Optimization Problems

by
James E. Falk

RESEARCH ANALYSIS CORPORATION
McLEAN, VIRGINIA

FOREWORD

This paper contains the mathematical validation of a Lagrangian technique for nonlinear programming that replaces the original problem by an auxiliary problem that is solvable by standard methods. The properties of the auxiliary problem are described and validated under certain conditions, and a number of applications are described. Several examples are presented in order to clarify various results.

Some parts of this paper have been extracted from the author's Ph.D. thesis in mathematics at the University of Michigan. The writing of this thesis was supported by National Science Foundation Grant GP-2215.

The author solicits criticisms, questions, and discussion of any of the conclusions reached.

Nicholas M. Smith
Head, Advanced Research Department

CONTENTS

Theory of Lagrange Multipliers for Constrained Optimization Problems

ABSTRACT

This paper treats an extension of one version of the classical Lagrange multiplier rule as applied to nonlinear programming problems. For a given problem, an auxiliary problem is defined and its properties are studied under various assumptions. In particular, when the given problem has a strictly convex objective function and concave constraints it is shown that the auxiliary problem is one of maximizing a concave differentiable function over an open set subject only to nonnegativity conditions. Some applications of this theory are presented along with the connection between the auxiliary problem and a "dual" of the given problem.

1. INTRODUCTION

Lagrange multipliers, in one form or another, have played an important role in the recent development of nonlinear programming theories. Indeed, perhaps the most important theoretical result in this field to date is the celebrated Kuhn-Tucker Theorem,[1] which is an extension of the classical Lagrange multiplier rule in its most common form (see Courant and Hilbert,[2] p 165). In the same paper, Kuhn and Tucker show the equivalence between convex programs and their associated "saddle value" problems.

Related to these concepts are the variations of the dual program formulated by Wolfe,[3] Huard,[4] and several others. This duality theory for nonlinear programming received impetus from its counterpart in linear programming, where it enjoys a very pleasing and useful symmetry. Early formulations of the dual of a nonlinear program did not enjoy perfect symmetry (for example, the dual of a convex program was not convex), and attempts to achieve it led to a closer study of the properties of Lagrangian functions (see Rockafellar[5] and Whinston[6] and their references).

A study of the Lagrangian function of a problem has proved useful from a computational standpoint. For example, Everett[7] has presented an interesting result that applies to general problems involving separable objective functions and constraints. The method essentially involves an iteration scheme in the space of Lagrange multipliers together with comparatively simple minimization operations at each iteration. Although it is clear how these minimization operations are to be performed, it is not clear how the optimal set of Lagrange multipliers are to be chosen.

Most of the work in this field has emphasized the best-known formulation of the Lagrange multiplier rule. There is another formulation (Ref 2, pp 231–32) based on the Legendre transformation that states the equivalence of a given equality constrained problem with a related but unconstrained optimization problem. The main purpose of this paper is to generalize this version of the Lagrange multiplier rule to handle inequality as well as equality constraints and to describe the structure of the related problem in some detail. It will, in fact, be shown that often a great deal of the structure of this related problem can be exploited computationally.

Section 2 contains the definitions of the various constituents of the related or auxiliary problem. These definitions can be made without reference to any particular hypothesis on the elements of the given problem, and some results may be obtained in this general setting.

In Sec 3 the discussion includes only convex programs with strictly convex objective functions. No differentiability assumptions are necessary. Although many of the results of this section hold for less restricted problems, the assumption of strict convexity seems to be the most concise and common hypothesis that can be made to ensure that the auxiliary problem is well behaved.

With these restrictions it will be shown that the auxiliary problem becomes one of maximizing a concave differentiable function over an open set subject only to nonnegativity conditions. This would appear to be a simple and useful procedure computationally, since any standard gradient-ascent technique could theoretically be employed on the auxiliary problem to obtain a solution of the given problem. Unfortunately the calculation of the gradient of the objective function of the auxiliary problem involves the solution of a nonlinear program, and, unless the given problem has a special structure, this solution may require an excessive amount of effort. On the other hand, many problems do have this special structure (e.g., separable programs) and for these problems the solution of the aforementioned nonlinear problem is easy. Lasdon,[8] Takahashi,[9] and Falk[10] have investigated such decomposable problems.

Takahashi views the auxiliary problem as the conjugate of a second related problem and uses known results of conjugate functions to verify his results. Although the theorems are stated correctly the proofs are incomplete since questions concerning the convexity of the domains of the functions involved are ignored.

In Sec 4 the auxiliary problem is related to the dual of the given problem as defined by Wolfe,[3] and it is shown that the two problems are essentially equivalent. This is important since the auxiliary problem is a convex program whereas the Wolfe dual generally is not.

Also in Sec 4 the theory is applied to decomposable and separable programs and to the problem of minimizing a quotient of two functions.

2. THE GENERAL CASE

The mathematical program to be discussed has the form
minimize

$$\{\phi(x) : f(x) \geqq 0, \; x \in C\} \tag{1}$$

where C is a subset of E^n and where $\phi : E^n \to E^1$ and $f : E^n \to E^m$. In general, for a given problem there are many ways to partition the constraining inequalities (or equalities), and hence the selection of a particular f and C in Eq 1 is somewhat arbitrary. Computational considerations discussed in Sec 4 indicate which constraints should be represented by f and which should be represented by C. It is assumed that $m \geq 1$.

Equality constraints have not been included explicitly in order to simplify later notation. Their inclusion would cause no theoretical problems as all the results that follow hold in their presence.[11]

The definitions that follow can be made without any additional hypothesis on ϕ, f, or C.

The Lagrangian function of Eq 1 is defined on $E^n \times E^m$ by the relation

$$\lambda(x, u) = \phi(x) - \langle u, f(x) \rangle . \tag{2}$$

A function γ is defined over its domain $D[\gamma]$ by means of the relations

$$D[\gamma] = \{u \geqq 0 : \gamma(\cdot, u) \text{ attains its minimum over } C\} \tag{3}$$

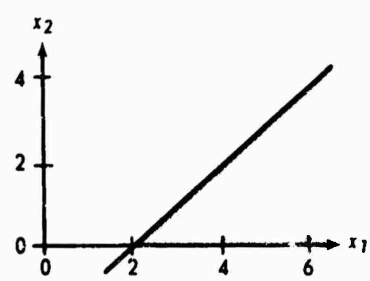
$$\gamma(u) = \min \{\lambda(x, u) : x \in C\}. \tag{4}$$

Thus u is in the domain of γ if and only if the function $\lambda(\cdot, u)$ attains a finite (absolute) minimum at some finite point x. The totality of points in C that minimize $\lambda(\cdot, u)$ for a given $u \in D[\gamma]$ is denoted by $X(u)$ and will be termed the "minimizing function." In general, $X(u)$ is a set function defined over $D[\gamma]$ into C. The function γ will be termed the "auxiliary function" of Eq 1, and the problem

$$\max \{\gamma(u) : u \in D[\gamma]\} \tag{5}$$

will be termed the "auxiliary problem" of Eq 1. These definitions may be illustrated by an example.

Example 1 (See Fig. 1)



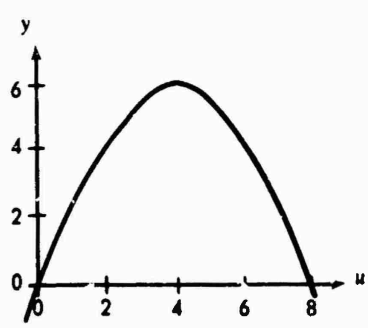


Fig. 1—Example 1

Minimize

$$\phi(x) = \frac{1}{2} x_1^2 + 2x_1 + \frac{1}{2} x_2^2 - x_2$$

subject to

$$f(x) = x_1 - x_2 - 2 \geq 0,$$

$$C : x_1, x_2 \geq 0,$$

$$\gamma(u) = \min_{x_1, x_2 \geq 0} \left\{\frac{1}{2} x_1^2 + 2x_1 + \frac{1}{2} x_2^2 - x_2 - u(x_1 - x_2 - 2)\right\}$$

$$= \min_{x_1 \geq 0} \left\{\frac{1}{2} x_1^2 + 2x_1 - u x_1\right\} + \min_{x_2 \geq 0} \left\{\frac{1}{2} x_2^2 - x_2 + u x_2\right\} + 2u.$$

Since each of these one-dimensional minimizations can be carried out for all values of $u \geq 0$ it follows that

$$D[\cdot] = (E^1)^+$$

Calculating the derivatives of the functions involved in this expression for γ and setting them equal to zero gives

$$X_1(u) = \begin{cases} 0 & 0 \leqq u \leqq 2 \\ u - 2 & u \geqq 2 \end{cases}$$

$$X_2(u) = \begin{cases} 0 & u \geqq 1 \\ 1 - u & 0 \leqq u \leqq 1. \end{cases}$$

Substituting these expressions into $\lambda(x,u)$ gives

$$\gamma(u) = \begin{cases} -\frac{1}{2}u^2 + 3u - \frac{1}{2} & 0 \leqq u \leqq 1 \\ 2u & 1 \leqq u \leqq 2 \\ -\frac{1}{2}u^2 + 4u - 2 & u \geqq 2. \end{cases}$$

If the equality constraints $g_i(x) = 0$ $(i = 1, \ldots, p)$ were added to problem 1 the auxiliary variables u_i $(i = 1, \ldots, p)$ associated with these constraints would not be restricted to be nonnegative in the definition of $D[\gamma]$.

It may be proved that if Eq 1 is the linear program

$$\min \{\langle c,x \rangle : Ax \geqq b,\ x \geqq 0\}$$

with $f(x) = Ax - b$ and $C = \{x : x \geq 0\}$, then Eq 5 is precisely its dual (see Falk.[11]) However, if some of the inequalities described by f are used to describe C, then problem 5 becomes a "piece-wise linear" program.

Example 2 (See Fig. 2)

Minimize

$$x_1 + 5x_2$$

subject to

$$f_1(x) = -2 + 2x_1 + x_2 \geqq 0$$

$$f_2(x) = -3 + x_1 + 3x_2 \geqq 0$$

$$C: \begin{cases} -3/2\,x_1 + x_2 \leqq 3 \\ x_1 - x_2 \leqq 2 \\ x_1, x_2 \geqq 0. \end{cases}$$

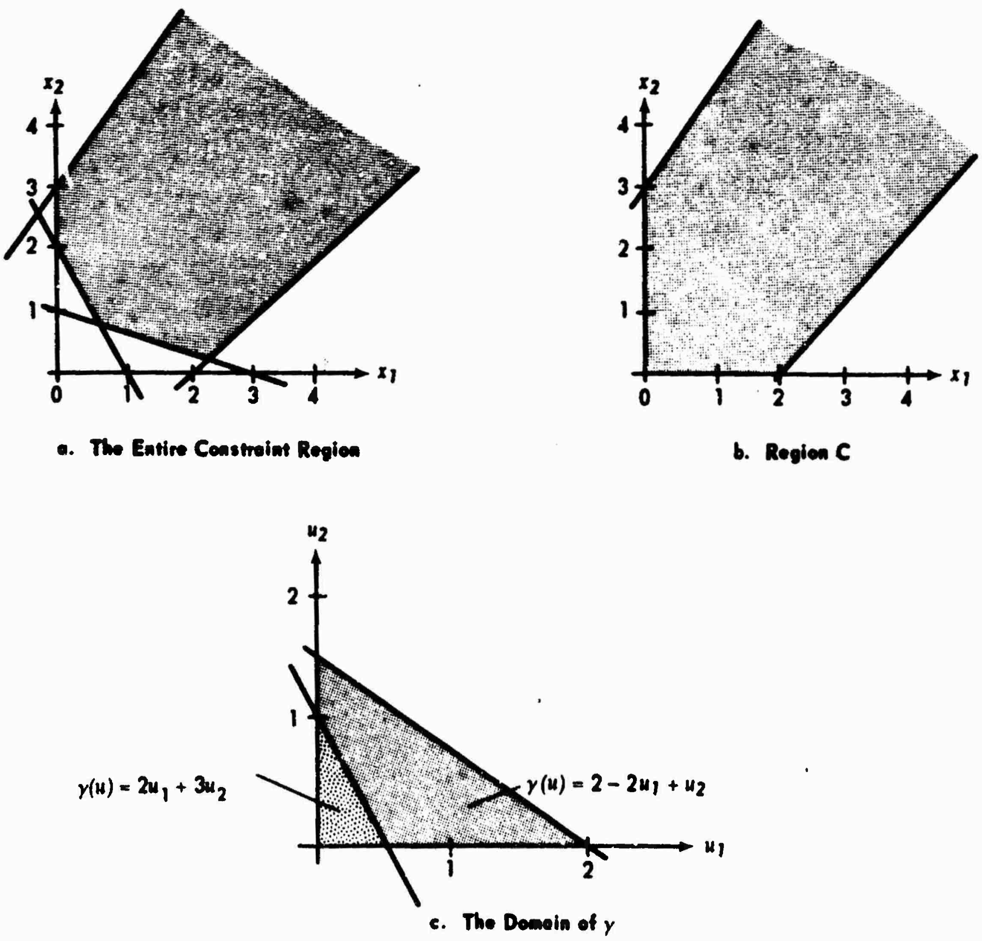
The Lagrangian function for this problem is

$$\begin{aligned}\lambda(x,u) &= x_1 + 5x_2 - u_1(-2 + 2x_1 + x_2) - u_2(-3 + x_1 + 3x_2)\\ &= (1 - 2u_1 - u_2)\,x_1 + (5 - u_1 - 3u_2)\,x_2 + 2u_1 + 3u_2\end{aligned}$$

and the corresponding auxiliary function becomes

$$\gamma(u) = \min_{x \in C} \{(1 - 2u_1 - u_2)\,x_1 + (5 - u_1 - 3u_2)\,x_2\} + 2u_1 + 3u_2.$$

The simplex tableau can be used to investigate the three vertices of C and to determine the corresponding sets of u_1 and u_2 for which these vertices are attained. The results are given in Fig. 2 along with the corresponding value of $\gamma(u)$.



a. The Entire Constraint Region

b. Region C

c. The Domain of γ

Fig. 2—A Linear Program with its Auxiliary Problem

If Eq 1 is the quadratic problem
minimize

$$\frac{1}{2} \langle x, Dx \rangle + \langle d, x \rangle$$

subject to

$$Ax \geqq b$$

where $f(x) = Ax - b$ and $C = E^n$ with D symmetric and positive definite then problem 5 becomes
maximize

$$-\frac{1}{2} \langle AD^{-1} A^T u, u \rangle + \langle AD^{-1} d + b, u \rangle - \frac{1}{2} \langle D^{-1} d, d \rangle$$

subject to

$$u \geqq 0.$$

This is the dual problem that Lemke[12] addresses in his method for quadratic programming.

Related to the function γ is the function $\bar{\gamma}$ defined by replacing "minimum" with "infimum" in definitions 3 and 4 and requiring that $u \in D[\bar{\gamma}]$ whenever $\lambda(\cdot, u)$ has a finite infimum over C. Clearly $D[\gamma] \subset D[\bar{\gamma}]$ and $\bar{\gamma}(u) = \gamma(u)$ for each $u \in D[\gamma]$. In many problems $D[\gamma] = D[\bar{\gamma}]$ (e.g., $D[\gamma] = D[\bar{\gamma}] = (E^m)^+$ if C is compact and if ϕ and f are continuous), although it is easy to find examples where the strict inclusion holds. One such example is constructed by setting

$$\phi(x) = e^{-x_2}$$

$$f(x) = x_1$$

$$C = E^2$$

Here

$$\gamma(u) = \min_{x_1, x_2} \{ e^{-x_2} - u x_1 \}$$

and this does not attain a minimum for any value of u so that $D[\gamma] = \phi$. However, for $u = 0$ the term e^{x_2} has an infimum of 0 so that $D[\bar{\gamma}] = \{0\}$.

This paper is primarily concerned with the function γ, although occasionally $\bar{\gamma}$ will be referred to in order to clarify certain results.

Theorem 1

The function γ is concave over convex subsets of $D[\gamma]$.

Proof: Fix $u^1, u^2 \in D[\gamma]$, $\alpha \geq 0$, $\beta \geq 0$, $\alpha + \beta = 1$, and assume that $u^3 = \alpha u^1 + \beta u^2 \in D[\gamma]$. Then

$$\gamma(u^3) = \min \{ \alpha\lambda(x, u^1) + \beta\lambda(x, u^2) : x \in C \}$$

$$\geq \alpha \min \{ \lambda(x, u^1) : x \in C \} + \beta \min \{ \lambda(x, u^2) : x \in C \}$$

$$= \alpha\gamma(u^1) + \beta\gamma(u^2)$$

and the proof is complete.

It is not true in general that $D[\gamma]$ is convex, even when Eq 1 is a convex program.

Example 3

Minimize

$$\{\max \{|x_1|, e^{-x_2}\} : x_1 = 0\}.$$

Each of the functions $\phi^1 : (x_1, x_2) \to |x_1|$ and $\phi^2 : (x_1, x_2) \to e^{-x_2}$ is convex. Since the maximum of two convex functions is convex, it follows that the function

$$\phi : (x_1, x_2) \to \max \{|x_1|, e^{-x_2}\}$$

is convex.

Pick

$$f(x) = x_1$$

$$C = E^2$$

Since the constraint is an equality, the value of u is unrestricted.

By definition

$$\gamma(u) = \min \{\max \{|x_1|, e^{-x_2}\} - ux_1\}$$

whenever this minimum exists. However, this minimum exists only when $u = 1$ and $u = -1$ so that

$$D\{\gamma\} = \{-1, 1\}$$

and is not convex. It is interesting to note that

$$X(1) = \{(x_1, x_2) : x_1 \geqq e^{-x_2}\}$$

and

$$X(-1) = \{(x_1, x_2) : x_1 \leqq -e^{-x_2}\}$$

are both unbounded sets. For convex problems such an example could not be constructed otherwise (see corollary to Theorem 8).

Note that $D[\gamma] = (E^m)^+$ when ϕ and f are continuous and when C is compact. The convexity of $D[\gamma]$ will be established for a different class of problems in Sec 3.

Since γ is concave over convex subsets of $D[\gamma]$, it is differentiable almost everywhere in int $D[\gamma]$. In order to calculate $\nabla\gamma = (\partial\gamma/\partial u_1, \ldots, \partial\gamma/\partial u_m)^T$ when it exists, it is necessary to establish a preliminary result.

Lemma 1

Let $u^* \in$ int $D[\gamma]$ and assume that the differential $d\gamma(u^*; .)$ exists. Let $g \in E^m$ be such that

$$\gamma(u^*) - \gamma(u) \geqq \langle g, u^* - u \rangle$$

for all $u \in D[\gamma]$. Then $\nabla\gamma(u^*) = g$.

Proof: Assume that $g \neq \nabla\gamma(u^*)$ and fix ϵ such that $0 < \epsilon < ||g - \nabla\gamma(u^*)||$. Since γ is differentiable at u^*, there is an open neighborhood of u^* of radius δ (denoted by $N(u^*; \delta)$) contained in $D[\gamma]$ such that

$$|\gamma(u^*) - \gamma(u) - \langle \nabla\gamma(u^*), u^* - u \rangle| < \epsilon||u - u^*||$$

for all $u \in N(u^*; \delta)$. Select $v^* \in N(u^*; \delta)$, $v^* \neq u^*$, on the ray emanating from u^* with direction $\nabla\gamma(u^*) - g$, i.e., select $\alpha > 0$ such that

$$v^* = u^* + \alpha(\nabla\gamma(u^*) - g) \in N(u^*; \delta).$$

With this selection of u^* it follows that

$$\begin{aligned}\langle \nabla\gamma(u^*) - g, v^* - u^* \rangle &= \alpha||\nabla\gamma(u^*) - g||^2 \\ &= ||v^* - u^*||\, ||\nabla\gamma(u^*) - g|| \\ &> \epsilon||v^* - u^*||\end{aligned}$$

so that

$$\langle g, u^* - v^* \rangle > \langle \nabla\gamma(u^*), u^* - v^* \rangle + \epsilon||v^* - u^*||.$$

The hypothesis of the theorem, together with this inequality, implies

$$\begin{aligned}\gamma(u^*) - \gamma(v^*) &\geq \langle g, u^* - v^* \rangle \\ &> \langle \nabla\gamma(u^*), u^* - v^* \rangle + \epsilon||v^* - u^*||,\end{aligned}$$

which violates Eq 5, and the proof is complete.

Theorem 2

Let $u^* \in \text{int } D[\gamma]$ and assume $d\gamma(u^*; \cdot)$ exists. Then

$$\nabla\gamma(u^*) = -f(x^*)$$

where x^* is any point in $X(u^*)$.

Proof:

$$\begin{aligned}\gamma(u) &= \min\{\phi(x^*) - \langle u, f(x) \rangle : x \in C\} \\ &\leq \phi(x^*) - \langle u, f(x^*) \rangle \\ &= \phi(x^*) - \langle u^*, f(x^*) \rangle - \langle u, f(x^*) \rangle + \langle u^*, f(x^*) \rangle \\ &= \gamma(u^*) + \langle u - u^*, -f(x^*) \rangle\end{aligned}$$

for all $u \in D[\gamma]$. Hence $-f(x^*)$ satisfies the condition of Lemma 1, and the proof is complete.

If γ attains its maximum at a point u^* where $d\gamma(u^*; \cdot)$ exists, the next result allows computation of the solution of problem 1.

Theorem 3

Assume that u^* maximizes γ and that γ is differentiable there. Then any point $x^* \in X(u^*)$ is a solution of problem 2. Furthermore, $\phi(x^*) = \gamma(u^*)$.

Proof: Since γ is maximized at u^* and is differentiable there it is necessary that

$$\nabla\gamma(u^*) = -f(x^*) = 0$$

where x^* is any point in $X(u^*)$. But

$$\begin{aligned}\gamma(u^*) &= \phi(x^*) - \langle u^*, f(x^*)\rangle \\ &\leq \phi(x) - \langle u^*, f(x)\rangle\end{aligned}$$

for all $x \in C$. For any feasible point x, $f(x) \geq 0$ so that $\langle u^*, f(x)\rangle \geq 0$. It follows that

$$\phi(x^*) \leq \phi(x)$$

for all feasible x, and the proof is complete.

Theorem 4

If x and u are feasible points of problem 1 and its auxiliary problem respectively, then

$$\gamma(u) \leq \phi(x).$$

Proof: Assume x and u are feasible. Then

$$\begin{aligned}\gamma(u) &= \min\{\phi(z) - \langle u, f(z)\rangle : z \in C\} \\ &\leq \phi(x) - \langle u, f(x)\rangle \\ &\leq \phi(x)\end{aligned}$$

and the proof is complete.

The following examples indicate the need to assume more structure on ϕ, f, and C in order to establish a close relation between problem 1 and its auxiliary problem. The objective functions of these problems are not convex.

Example 4 (See Fig. 3)

Minimize

$$\{-x^2 : 1 - 2x = 0, 0 \leq x \leq 1\}.$$

Choosing $f(x) = 1 - 2x$ and $C = \{x : 0 \leq x \leq 1\}$ it is found that $D[\gamma] = E^1$ and γ attains its maximum at $u^* = 1/2$. However $x^* = 1/2 \notin X(u^*) = \{0,1\}$ and $\gamma(u^*) < \phi(x^*)$. Hence both problems are feasible and have optimal solutions, but these solutions are not directly related.

Example 5 (See Fig. 3)

Minimize

$$\{-x^2 : 1 - 2x = 0, x = 0 \text{ or } 1\}.$$

Choosing f as above and $C = \{0,1\}$ it is easily seen that the auxiliary problem is feasible and has an optimal solution whereas the stated problem is not feasible.

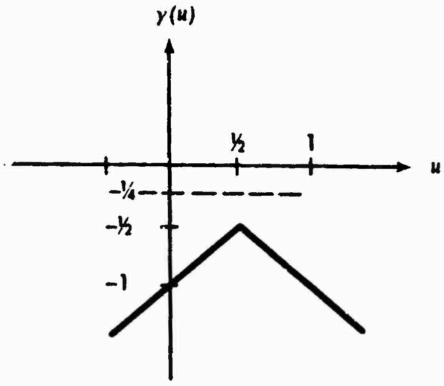


Fig. 3—The Auxiliary Function

Example 6

Minimize

$$\{x^3 : x \geq 0\}.$$

Setting $f(x) = x$ and $C = E^1$ it follows that $D[\gamma]$ is empty so that the auxiliary problem is not feasible, whereas the stated problem is feasible and has an optimal solution.

The following principle has applications to decomposable and separable programming which will be pointed out in more detail in Sec 4. The theorem is stated here because no special hypotheses are needed on $\phi, f,$ and C.

Suppose that problem 1 has the form

minimize

$$\{\phi(x) + \Psi(y) : f(x) + g(y) \geq 0,\ x \in C,\ y \in D\}. \tag{6}$$

Let γ denote the auxiliary function of this problem and (X, Y) its minimizing function. Let γ' and γ'' denote the auxiliary functions of the two problems

minimize

$$\{\phi(x) : f(x) \geq 0,\ x \in C\} \text{ and} \tag{7}$$

minimize

$$\{\Psi(y) : g(y) \geq 0,\ y \in D\}. \tag{8}$$

Let X' and Y' denote the minimizing functions of Eqs 7 and 8 respectively.

Theorem 5

Problems 6, 7, and 8 are related in the following manner:

(a) $$\gamma = \gamma' + \gamma''$$

(b) $$D[\gamma] = D[\gamma'] \cap D[\gamma'']$$

(c) $$(X, Y) = (X', Y')$$

Proof: The theorem is a direct consequence of the relation

$$\begin{aligned}\gamma(u) &= \min\{\phi(x) + \Psi(y) - \langle u, f(x) + g(y)\rangle : x \in C,\ y \in D\} \\ &= \min\{\phi(x) - \langle u, f(x)\rangle : x \in C\} \\ &\quad + \min\{\Psi(y) - \langle u, g(y)\rangle : y \in D\}.\end{aligned}$$

The next theorem states that any feasible point u of the auxiliary problem yields, via $X(u)$, a solution of another problem related to problem 1.

__Theorem 6__ (Everett)[7]

If $u^* \in D[\gamma]$ then a point x^* in $X(u^*)$ is a solution of the problem minimize

$$\{\phi(x) : f(x) \geqq f(x^*),\ x \in C\}.$$

Proof: By definition, $x^* \in C$ and

$$\gamma(u^*) = \phi(x^*) - \langle u^*, f(x^*)\rangle \leqq \phi(x) - \langle u^*, f(x)\rangle$$

for all $x \in C$. Since $u^* \geqq 0$ it follows that

$$\begin{aligned}\phi(x^*) &\leqq \phi(x) - \langle u^*, f(x) - f(x^*)\rangle \\ &\leqq \phi(x)\end{aligned}$$

for all $x \in C$ and $f(x) \geqq f(x^*)$.

3. THE STRICTLY CONVEX CASE

Unless otherwise stated, throughout this section it is assumed that the Lagrangian function $\lambda(\cdot, u)$ defined in problem 2 is strictly convex for each $u \in D[\gamma]$. Such would be the case, for example, if ϕ is strictly convex and each f_i is concave. No differentiability assumptions are required. The set C is assumed to be closed and convex but not necessarily compact.

__Theorem 7__

If $\lambda(\cdot, u)$ is strictly convex for each $u \in D[\gamma]$ and C is closed and convex, then $D[\gamma]$ is an open set relative to $(E^m)^+$.

Proof: Fix $u^* \in D[\gamma]$ and let $x^* = X(u^*)$. [Since $\lambda(\cdot, u^*)$ is strictly convex and has a minimum over C, it must have a unique minimum.] Let $N(x^*; \epsilon)$ be a neighborhood of x^* of radius ϵ where $\epsilon > 0$ and $C \cap N(x^*; \epsilon) \neq \phi$. If such an ϵ cannot be found, then C consists of the single point x^* and the theorem is trivial. Let

$$\mu_1 = \lambda(x^*, u^*)$$
$$\mu_2 = \min\{\lambda(x, u^*) : x \in C \cap \partial N(x^*, \epsilon)\}$$

(The symbol ∂ denotes the boundary operator.)
Then $\mu_2 > \mu_1$ because $\lambda(\cdot, u^*)$ is strictly convex as a function of x. Let

$$\nu_1 = \|f(x^*)\|$$
$$\nu_2 = \max\{\|f(x)\| : x \in C \cap \partial N(x^*; \epsilon)\}$$

and set

$$0 < \delta < \begin{cases} 1 & \text{if } \nu_1 = \nu_2 = 0 \\ (\mu_2 - \mu_1)/3\nu_1 & \text{if } \nu_1 \neq 0, \nu_2 = 0 \\ (\mu_2 - \mu_1)/3\nu_2 & \text{if } \nu_1 = 0, \nu_2 \neq 0 \\ \min\{(\mu_2 - \mu_1)/3\nu_1, & \text{if } \nu_1 \neq 0, \nu_2 \neq 0 \\ (\mu_2 - \mu_1)/3\nu_2\}. & \end{cases}$$

Then, if $\|u - u^*\| < \delta$, $u \in D[\gamma]$, it follows that

$$\begin{aligned} &|\{\phi(x^*) - \langle u, f(x^*)\rangle\} - \{\phi(x^*) - \langle u^*, f(x^*)\rangle\}| \\ &= |\langle u^* - u, f(x^*)\rangle| \\ &\leqq \|u^* - u\| \, \|f(x^*)\| \\ &= \nu_1 \|u - u^*\| \\ &< (\mu_2 - \mu_1)/3 \end{aligned}$$

so that $\lambda(\cdot, u)$ evaluated at x^* differs from μ_1 by less than $(\mu_2 - \mu_1)/3$.

Furthermore, if $x' \in C \cap \partial N(x^*; \epsilon)$ it follows that

$$\begin{aligned} |\lambda(x', u^*) - \lambda(x', u)| &= |\langle u - u^*, f(x')\rangle| \\ &\leqq \|u - u^*\| \, \|f(x')\| \\ &\leqq \nu_2 \|u - u^*\| \\ &< (\mu_2 - \mu_1)/3 \end{aligned}$$

so that

$$\lambda(x', u^*) < (\mu_2 - \mu_1)/3 + \lambda(x', u).$$

Hence

$$\mu_2 - (\mu_2 - \mu_1)/3 < \lambda(x', u).$$

Now $\lambda(\cdot, u)$ must be minimized somewhere in $C \cap \bar{N}(x^*; \epsilon)$ since this set is compact. Its minimum cannot occur on $C \cap \partial N(x^*; \epsilon)$, since there is a point x^* in $C \cap N(x^*; \epsilon)$ giving a lower value to $\lambda(\cdot, u)$ than any point on $C \cap \partial N(x^*; \epsilon)$. Since ϕ is strictly convex, the minimum of $\lambda(\cdot, u)$ is the unique global minimum of $\lambda(\cdot, u)$ over C, and the proof is complete.

The strict convexity is essential in this proof since, for example, a linear program has a closed convex polyhedron for the domain of its auxiliary function.

The next theorem, together with Theorem 1, shows that the auxiliary problem is a convex program.

<u>Theorem 8</u>

If $\lambda(\cdot, u)$ is strictly convex for each $u \in D[\gamma]$ and C is closed and convex, then $D[\gamma]$ is convex.

Proof: Fix u^1 and u^2 in $D[\gamma]$ and let $u^3 = \alpha u^1 + \beta u^2$ where $\alpha, \beta \geq 0$, $\alpha + \beta = 1$. Since

$$\inf \{\lambda(x,u^3) : x \in C\} \geq \alpha \inf \{\lambda(x,u^1) : x \in C\} + \beta \inf \{\lambda(x,u^2) : x \in C\}$$

it follows that $\lambda(\cdot,u^3)$ is bounded below on C. It must be shown that the greatest lower bound of $\lambda(\cdot,u^3)$ over C is actually attained at some point of C.

Let

$$\mu_i = \inf \{\lambda(x,u^i) : x \in C\} \quad (i = 1, 2, 3)$$

and set

$$\mu_4 > \max \{\mu_1, \mu_2, \mu_3\}.$$

Consider the sets

$$L^i(\mu_4) = \{x \in C : \lambda(x,u^i) \leq \mu_4\} \quad (i = 1, 2, 3).$$

Since $\lambda(\cdot,u^1)$ and $\lambda(\cdot, u^2)$ have unique minima in C, it follows that the sets $L^1(\mu_4)$ and $L^2(\mu_4)$ are bounded sets because all nonempty level sets of a convex function are bounded in the same directions.[13] But

$$L^3(\mu_4) \subset L^1(\mu_4) \cup L^2(\mu_4)$$

so that $L^3(\mu_4)$ is also bounded and hence compact. Thus $\lambda(\cdot,u^3)$ attains its minimum in C, and the proof is complete.

Although the strict convexity of ϕ is a sufficient condition for the convexity of $D[\gamma]$, it is not necessary since $D[\gamma]$ is convex in the linear programming case. Example 3 following Theorem 1 illustrates the need for some condition such as the strict convexity of $\lambda(\cdot,u)$ to ensure the convexity of $D[\gamma]$. The corollary that follows relaxes the strict convexity assumption somewhat.

The "relative interior" of $D[\gamma]$ refers to the interior of $D[\gamma]$ with respect to the smallest linear manifold containing $D[\gamma]$.

<u>Corollary</u>

If $\lambda(\cdot,u)$ is convex and C is closed and convex (so that problem 1 is a convex program), and if $X(u^1)$ is bounded for some $u^1 \in D[\gamma]$ then the relative interior of $D[\gamma]$ is convex.

Proof: Let u^3 be any point in the relative interior of $D[\gamma]$. Then

$$u^3 = \alpha u^1 + \beta u^2 \quad (\alpha, \beta > 0,\ \alpha + \beta = 1)$$

where $u^2 \in D[\gamma]$ is on the ray emanating from u^1 and passing through u^3. Using the notation of the above theorem we obtain as before

$$\mu^3 \geq \alpha\mu^1 + \beta u^2.$$

Let $y^* \in X(u^2)$. Then $\lambda(x,u^2) \geq \lambda(y^*,u^2)$ for all $x \in C$; hence

$$
\begin{aligned}
L_3(\mu^3) &= \{x \in C : \alpha\lambda(x,u^1) + \beta\lambda(x,u^2) \leqq \lambda(y^*,u^3)\} \\
&= \{x \in C : \alpha\lambda(x,u^1) \leqq \alpha\lambda(y^*,u^1) + \beta[\lambda(y^*,u^2) - \lambda(x,u^2)]\} \\
&\subset \{x \in C : \lambda(x,u^1) \leqq \alpha\lambda(y^*,u^1)\} \\
&= L_1(\lambda(y^*,u^1)) \text{ which is a bounded set.}
\end{aligned}
$$

It is not empty since $y^* \in L_3(\lambda(y^*,u^3))$.

Hence $\lambda(\cdot,u^3)$ attains a minimum over C. Since u^3 was an arbitrary point in the relative interior of $D[\gamma]$, the proof is complete.

The next theorem categorizes the minimizing function X in the strictly convex case. Note that $X(u)$ consists of a single point for each u and hence may be considered a function in the usual sense.

<u>Theorem 9</u>

If $\lambda(\cdot,u)$ is strictly convex for each $u \in D[\gamma]$ and C is closed and convex, then X is a continuous function on $D[\gamma]$.

Proof: Fix $u^* \in D[\gamma]$ and $\epsilon > 0$. It must be shown that there is a $\delta > 0$ such that $\|u - u^*\| < \delta$, $u \in D[\gamma]$ implies that $\|X(u) - X(u^*)\| < \epsilon$. Set $x^* = X(u^*)$ and

$$M > \max \{\|f(x) - f(x^*)\| : x \in C \cap \partial N(x^*;\epsilon)\}.$$

(If $C \cap \partial N(x^*;\epsilon)$ is empty the proof is immediate.) Let $\delta > 0$ be any number such that

$$(1/M)\,|\phi(x) - \phi(x^*) - \langle u^*, f(x) - f(x^*)\rangle| > \delta$$

for all $x \in C \cap \partial N(x^*;\epsilon)$.

If $u \in D[\gamma] \cap N(u^*;\delta)$ then

$$
\begin{aligned}
M\delta &> \|f(x) - f(x^*)\|\,\|u - u^*\| \\
&\geqq |\langle f(x) - f(x^*), u - u^*\rangle| \\
&\geqq \langle f(x) - f(x^*), u - u^*\rangle.
\end{aligned}
$$

But

$$|\phi(x) - \langle u^*, f(x)\rangle| - |\phi(x^*) - \langle u^*, f(x^*)\rangle| > M\delta$$

for all $x \in C \cap \partial N(x^*;\epsilon)$

so that

$$\phi(x) - \langle u, f(x)\rangle > \phi(x^*) - \langle u, f(x^*)\rangle$$

for all $x \in C \cap \partial N(x^*;\epsilon)$.

Since $u \in D[\gamma]$, $\lambda(\cdot,u)$ has a minimum over C, and this last inequality shows that this minimum cannot occur on $C \cap \partial N(x^*;\epsilon)$. The strict convexity of $\lambda(\cdot,u)$ requires that it be minimized in $C \cap N(x^*;\epsilon)$, and the proof is complete.

Theorem 10†

If $\lambda(\cdot, u)$ is strictly convex for each $u \in D[\gamma]$ and C is closed and convex, then the partial derivatives $\partial\gamma/\partial u_i$ exist and are continuous throughout int $D[\gamma]$, and hence γ is differentiable there. Moreover, if $u^* \in D[\gamma]$ and $u_i^* = 0$ then the right-hand partial $\partial\gamma/\partial u_i^+$ exists at u^*.

Proof: Fix $u^* \in D[\gamma]$ and $h > 0$. By letting e^i denote the ith unit vector

$$\frac{\gamma(u^* + he^i) - \gamma(u^*)}{h} \leq \frac{1}{h}\{\lambda(X(u^*), u^* + he^i) - \lambda(X(u^*), u^*)\}$$
$$= -f_i(X(u^*)).$$

On the other hand

$$\frac{\gamma(u^* + he^i) - \gamma(u^*)}{h} \geq \frac{1}{h}\{\lambda(X(u^* + he^i), u^* + he^i) - \lambda(X(u^* + he^i, u^*))\}$$
$$= -f_i(X(u^* + he^i)).$$

Since f_i and X are continuous, it follows that

$$\lim_{h\to 0}\{-f_i(X(u^* + he^i))\} = -f_i(X(u^*)) = \frac{\partial\gamma}{\partial u_i^+}\bigg|u^*.$$

For $u_i > 0$ a similar proof shows that $\partial\gamma/\partial u_i|u^* = -f_i(X(u^*))$ and the proof is complete.

This theorem, together with Theorems 7 and 8, allows one to find the solution of the auxiliary problem by employing any standard gradient-ascent technique that takes into account the nonnegativity condition on u (this last restriction is unnecessary if only equality constraints are present on the original problem).

Although γ is continuously differentiable in int $D[\gamma]$ it is not, in general, twice continuously differentiable. For example, when the objective function and the constraining functions are separable and when $C = (E^n)^+$, the region $D[\gamma]$ is partitioned into several subregions by n hyperplanes (see Ref 11, p 92). The degree of differentiability of γ inside these subregions depends primarily on the degree of differentiability of ϕ and the f_i. Typically γ is not twice continuously differentiable on the common boundaries of these subregions.

Example 7 (See Fig. 4)

Minimize

$$\phi(x) = \frac{1}{2}x_1^2 + 2x_1 + e^{-x_2} + 3x_2 + \frac{1}{2}x_3^2 - 4x_3$$

subject to

$$x_1 - x_3 = 1$$
$$x_1 + x_2 = 2$$
$$x_1, x_2, x_3 \geq 0.$$

†The author is indebted to G. P. McCormick for suggesting the proof presented here for this theorem.

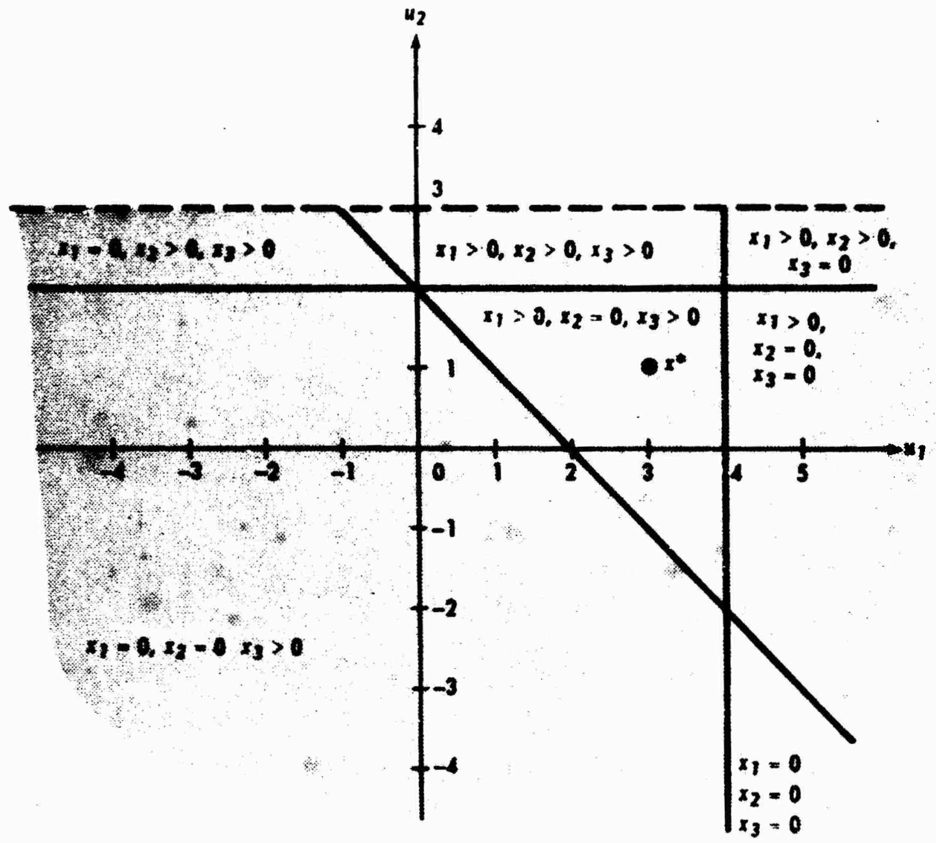


Fig. 4—The Domain of γ

The auxiliary function of this problem is defined by

$$\gamma(u) = \min_{x_1, x_2, x_3 \geq 0} \left\{ \frac{1}{2} x_1^2 + 2x_1 + e^{-x_2} + 3x_2 + \frac{1}{2} x_3^2 - 4x_3 \right.$$
$$\left. - u_1(x_1 - x_3 - 1) - u_2(x_1 + x_2 - 1) \right\}.$$

Because of the separability of $\lambda(x, u)$ in the x variables, it may be written that

$$\gamma(u) = \min_{x_1 \geq 0} \left\{ \frac{1}{2} x_1^2 + 2x_1 - (u_1 + u_2) x_1 \right\}$$
$$+ \min_{x_2 \geq 0} \left\{ e^{-x_2} + 3x_2 - u_2 x_2 \right\}$$
$$+ \min_{x_2 \geq 0} \left\{ \frac{1}{2} x_3^2 - 4x_3 + u_1 x_3 \right\}$$
$$+ u_1 + u_2.$$

Each of these minimizations is easily performed and it is found that

$$x_1(u) = \begin{cases} u_1 + u_2 - 2 & \text{when } u_1 + u_2 \geqq 2 \\ 0 & \text{otherwise} \end{cases}$$

$$x_2(u) \geq \begin{cases} \text{undefined} & \text{when } u_2 > 3 \\ -\ln(3 - u_2) & 2 \leqq u_2 < 3 \\ 0 & u_2 \leqq 2 \end{cases}$$

$$x_3(u) \geq \begin{cases} 4 - u_1 & \text{when } u_1 \leqq 4 \\ 0 & \text{otherwise.} \end{cases}$$

The domain of γ is given by the inequalities involved on the definitions of the $x_i(u)$'s above. Figure 4 illustrates $D[\gamma]$ along with the subregions defined by these inequalities. Note that

$$\nabla\gamma(3,1) = b - Ax(3,1)$$

$$= \begin{pmatrix} 1 \\ 2 \end{pmatrix} - \begin{pmatrix} 1 & 0 & -1 \\ 1 & 1 & 0 \end{pmatrix}\begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

so that the point $x^* = (3,1)$ maximizes γ. This point could have been found using a standard gradient-ascent technique that takes into consideration the possible bounds on $D[\gamma]$. In many cases $D[\gamma] = E^m$ so that no special care is necessary.

The following two theorems state that problem 1 and its auxiliary problem are equivalent in the sense that the solution of one provides a solution of the other and that their optimal values are equal. The proof of Theorem 11 follows directly from Theorems 7 and 10.

Theorem 11

If γ is maximized over $D[\gamma]$ at u^* then $x^* = X(u^*)$ is the solution of problem 1. Furthermore, $\gamma(u^*) = \phi(x^*)$.

Proof: Since γ attains its maximum at u^* it is necessary that

$$\frac{\partial\gamma}{\partial u_i}\Big| u^* = 0 \quad \text{if } u_i^* > 0$$

$$\frac{\partial\gamma}{\partial u_i^+}\Big| u^* \leqq 0 \quad \text{if } u_i^* = 0$$

i.e.,

$$f_i(x^*) = 0 \quad \text{if } u_i^* > 0$$

$$f_i(x^*) \geqq 0 \quad \text{if } u_i^* = 0$$

so that x^* is feasible. Moreover

$$\gamma(u^*) = \phi(x^*) - \langle u^*, f(x^*)\rangle = \phi(x^*)$$

so that ϕ attains its minimum at x^* by Theorem 4, and the proof is complete.

To prove the converse of Theorem 11 we may modify the proof of a similar theorem found in Arrow et al.[14] It is necessary to make the additional assumption that there is a point $x^0 \in C$ such that $f(x^0) > 0$ (which im-

plies that $\{x \mid f(x) \geqq 0\}$ has a nonempty interior relative to C). This is a common assumption that is often employed when dealing with concave inequalities (see Arrow et al., p 34).[14] The assumption of strict convexity may be dropped for this proof.

Theorem 12

Let $\phi, -f_1, \ldots, -f_m$ be convex functions defined over C and assume that there is a point $x^0 \in C$ such that $f(x^0) > 0$. If problem 1 has a solution x^*, then its auxiliary problem is feasible and has a solution u^*. Moreover $\phi(x^*) = \gamma(u^*)$.

Proof: Define two sets T and T′ in E^{m+1} by

$$T = \{(\tau, t) : \tau \geqq \phi(x),\ f(x) \geqq t \text{ for some } x \in C\}$$

$$T' = \{(\tau', t') : \phi(x^*) > \tau',\ t' \geqq 0\}.$$

It is easily seen that $T \cap T' = \phi$. Since T and T′ are convex, there is a hyperplane that separates them; i.e., there is an $n + 1$ vector $(\nu^*, v^*) \neq 0$ such that $\nu^*\tau + \langle v^*, t \rangle \geq \nu^*\tau' + \langle v^*, t' \rangle$ for all $(\tau, t) \in T$ and $(\tau', t') \in T'$.

It will now be shown that

$$\nu^* \geqq 0$$

$$v^* \leqq 0.$$

Fix t, τ', and t' in the above inequality and let $\tau \to \infty$. If $\nu^* < 0$ the inequality would become violated for sufficiently large τ. A similar argument fixing τ, t, and τ' and allowing t' to become arbitrarily large yields $v^* \leqq 0$. Moreover, $\nu^* \neq 0$ because if $\nu^* = 0$ then

$$\langle v^*, f(x^0) \rangle \geqq 0$$

since $(\phi(x^0), f(x^0)) \in T$, and $(\phi(x^*) - 1, 0) \in T'$. But $v^* \leqq 0$, $v^* \neq 0$ ($(\nu^*, v^*) \neq 0$), and $f(x^0) > 0$ so that

$$\langle v^*, f(x^0) \rangle < 0,$$

which is a contradiction.

Set

$$u^* = -\left(\frac{1}{\nu^*}\right) v^*.$$

Thus

$$\tau - \langle u^*, t \rangle \geqq \tau' - \langle u^*, t' \rangle.$$

Since $(\phi(x), f(x)) \in T$ for any feasible x, it follows that

$$\phi(x) - \langle u^*, f(x) \rangle \geqq \tau' - \langle u^*, t' \rangle.$$

Setting $t' = 0$ and taking the supremum of the right-hand side gives

$$\phi(x) - \langle u^*, f(x) \rangle \geqq \phi(x^*),$$

which implies immediately the conclusion of the theorem.

4. SOME APPLICATIONS

In this section some applications of the theory developed in Secs 2 and 3 are briefly described. A paper describing applications in more detail is in preparation. Lasdon[8] and Takahashi[9] contain additional applications to resource and multistage allocation problems.

Duality

The importance of duality theory in linear programming has led to the concept of the dual of a nonlinear program. The formulation in this section is from P. Wolfe.[3]

The problem that Wolfe considers has the form

minimize

$$\phi(x)$$

subject to

$$f(x) \geqq 0 \tag{9}$$

where ϕ and the m components of $-f$ are convex and continuously differentiable functions.

The Wolfe dual of Eq 9 has the form

maximize

$$\Psi(x,u) = \phi(x) - \langle u, f(x) \rangle$$

subject to

$$\begin{aligned} \nabla_x \Psi(x,u) &= 0 \\ u &\geqq 0. \end{aligned} \tag{10}$$

The feasible set of Eq 10 is denoted by $D[\Psi]$. Note that Eq 10 contains the variable x as well as the dual variable u. In general Eq 10 does not describe a convex set.

Theorem 13

The auxiliary problem of Eq 9 is equivalent to Eq 10 in the following sense:

(a) $$D[\Psi] = \{(x,u) : u \in D[\gamma],\ x \in X(u)\}$$

(b) $$\Psi(x,u) = \gamma(u) \text{ for each } (x,u) \in D[\Psi].$$

Proof: Since ϕ and $-f_i$ ($i = 1, \ldots, m$) are convex, a necessary and sufficient condition that $\Psi(\cdot, u)$ be minimized over E^n is that

$$\nabla_x \Psi(x,u) = 0.$$

Hence, if $(x, u) \in D[\Psi]$, then x minimizes $\Psi(\cdot, u)$ over E^n and conversely. This proves statement a. Statement b is immediate from the definition of γ.

While Eq 10 is not a convex program, the auxiliary problem of Eq 9 is, at least when ϕ is strictly convex. Also the primal variable x does not occur in the auxiliary problem since it has been replaced by $X(u)$.

Decomposable and Separable Programming

Suppose Eq 1 has the form
minimize

$$\phi(x) = \sum_{i=1}^{p} \phi_i(x_i)$$

subject to

$$f(x) = \sum_{i=1}^{p} f^i(x_i) \geq 0$$

$$g_i(x_i) \geq 0 \quad (i = 1, \ldots, p)$$

where ϕ is strictly convex, each f^i and g_i are concave, $x = (x_1, \ldots, x_p)^T$, and x_i is a vector having n_i components. Such a problem is said to be decomposable and if $n_i = 1$ $(i = 1, \ldots, p)$ it is said to be separable (completely decomposable). Letting $C = \{x : g_i(x_i) \geqq 0\}$, by Theorem 5, yields

$$\gamma(u) = \sum_{i=1}^{p} \min \{\phi_i(x^i) - \langle u, f^i(x_i)\rangle : g_i(x_i) \geq 0\}.$$

Hence, if a gradient-ascent procedure is used to maximize γ, the essential quantities $\gamma(u)$ and $\nabla\gamma(u)$ can be obtained by solving p nonlinear programs for each u. The ith program involves n_i variables. Thus the solution of a decomposable program is obtained by solving p smaller subprograms for a sequence of u's tending to u^*. In the separable case the p subprograms involve a single variable only and, in many cases, $X_i(u)$ can be expressed analytically. In the important special case where f has a single component much more can be said about the solution of Eq 1 (see Falk[11]).

Minimizing Quotients

Suppose one is seeking the solution of a problem having the form
minimize

$$\phi(x)/\Psi(x)$$

subject to

$$x \in S \tag{11}$$

where ϕ and $-\Psi$ are convex, S is closed and convex, and $\Psi(x) > 0$ for all $x \in S$. It will be assumed that ϕ is strictly convex.

The function δ may be defined over its domain $D[\delta]$ by the relations

$$D[\delta] = \{\mu \geqq 0 : \phi(\cdot) - \mu\Psi(\cdot) \text{ has a minimum over } S\}$$

$$\delta(\mu) = \min\{\phi(x) - \mu\Psi(x) : x \in S\}$$

By the theorems of Sec 3 it is known that $D[\delta]$ is open with respect to $(E^1)^+$ and convex, and δ is concave.

Theorem 14

δ is a monotone decreasing function.

Proof: Let $\mu' \leq \mu''$ where $\mu', \mu'' \in D[\delta]$. Since $\Psi(x) > 0$ for $x \in S$ we have

$$\phi(x) - \mu'\Psi(x) \geq \phi(x) - \mu''\Psi(x)$$

so that

$$\delta(\mu') \geq \phi(X(\mu')) - \mu''\Psi(X(\mu'))$$

where $X(\mu')$ minimizes $\phi(\cdot) - \mu'\Psi(\cdot)$ over S. Hence

$$\delta(\mu') \geq \delta(\mu'')$$

and the proof is complete.

Since δ is continuous and monotone decreasing, the set of points μ^* for which $\delta(\mu^*) = 0$ is connected and compact. The next theorem characterizes the points of this set.

Theorem 15

There is a point μ^* such that $\delta(\mu^*) = 0$ if and only if μ^* is the optimum value of the objective function of Eq 11. Moreover $X(\mu^*)$ is a solution of Eq 11.

Proof: If $\delta(\mu^*) = 0$ it follows that

$$\phi(x) - \mu^*\Psi(x) \geq \delta(\mu^*) = 0 \text{ for } x \in S.$$

Hence

$$\phi(x)/\Psi(x) \geq \mu^* \text{ for all } x \in S.$$

Moreover

$$\phi(X(\mu^*)) - \mu^*\Psi(X(\mu^*)) = 0$$

so that

$$\phi(X(\mu^*))/\Psi(X(\mu^*)) = \mu^*.$$

The other half of the proof is similar.

Hence the problem of minimizing the quotient of two functions can be viewed as a sequence of minimization problems not involving quotients. In many cases, each problem in this sequence of problems may require a minimum of computational effort compared to the original problem. The sequence of problems to be solved is formed sequentially in a manner that will locate a zero of the concave decreasing function δ.

REFERENCES

CITED REFERENCES

1. H. W. Kuhn and A. W. Tucker, "Non-Linear Programming," in Proceedings of the Second Berkeley Symposium of Mathematical Statistics and Probability, J. Neyman (ed), University of California Press, Berkeley, Calif., 1951.
2. R. Courant and D. Hilbert, Methods of Mathematical Physics, Vol 1, Interscience Publishers, Inc. (now John Wiley & Sons, Inc., New York), 1953.
3. P. Wolfe, "A Duality Theorem for Nonlinear Programming," Quart. Appl. Math., 19: (Oct 61).
4. P. Huard, "Dual Programs," in Recent Advances in Mathematical Programming, R. L. Graves and P. Wolfe (ed), McGraw-Hill Book Company, New York, 1963.
5. R. T. Rockafellar, "Convex Functions and Dual Extremum Problems," doctoral dissertation, Harvard University, Cambridge, Mass., 1963.
6. A. Whinston, "Conjugate Functions and Symmetric Duality," Cowles Foundation Discussion Paper, 162: (Nov 63).
7. H. Everett, "Generalized LaGrange Multiplier Method for Solving Problems of Optimal Allocation of Resources," Opns. Res. 11(3): (May–Jun 63).
8. Leon S. Lasdon, "A Multi-Level Technique for Optimization," doctoral dissertation, Case Institute of Technology, Cleveland, Ohio, 1964.
9. I. Takahashi, "Variable Separation Principle for Mathematical Programming," J. Opns. Res. Soc. Japan, 6 (2): (Feb 64).
10. J. E. Falk, "An Algorithm for Separable Convex Programming under Linear Equality Constraints," RAC-TP-148, Research Analysis Corporation, Mar 65.
11. ——, "A Constrained Lagrangian Approach to Nonlinear Programming," doctoral dissertation, University of Michigan, May 65.
12. C. E. Lemke, "A Method of Solution for Quadratic Programming," Mgt. Sci. 8 (4): (Jul 62).
13. W. Fenchel, Convex Cones, Sets and Functions, Princeton University Press, Princeton, N. J., 1953.
14. Kenneth J. Arrow, L. Hurwicz, and H. Uzawa, Studies in Linear and Non-Linear Programming, Stanford University Press, Stanford, Calif., 1958.

ADDITIONAL REFERENCES

S. E. Dreyfus, Dynamic Programming and the Calculus of Variations, Academic Press Inc., New York, 1965.

A. V. Fiacco and G. P. McCormick, "The Sequential Unconstrained Minimization Technique for Nonlinear Programming, A Primal-Dual Method," Mgt. Sci. 10 (2): (Jan 64).

Unclassified
Security Classification

DOCUMENT CONTROL DATA - R&D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) Research Analysis Corporation McLean, Virginia 22101	2a. REPORT SECURITY CLASSIFICATION Unclassified
	2b. GROUP None

3. REPORT TITLE
THEORY OF LAGRANGE MULTIPLIERS FOR CONSTRAINED OPTIMIZATION PROBLEMS

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)
Technical Paper

5. AUTHOR(S) (First name, middle initial, last name)
James E. Falk

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
July 1967	29	16

8a. CONTRACT OR GRANT NO. DA 44-188-ARO-1	9a. ORIGINATOR'S REPORT NUMBER(S) RAC-TP-262
b. PROJECT NO. 007.116	
c.	9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)
d.	

10. DISTRIBUTION STATEMENT

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY

13. ABSTRACT

This paper treats an extension of one version of the classical Lagrange multiplier rule as applied to nonlinear programming problems. For a given problem, an auxiliary problem is defined and its properties are studied under various assumptions. In particular, when the given problem has a strictly convex objective function and concave constraints it is shown that the auxiliary problem is one of maximizing a concave differentiable function over an open set subject only to nonnegativity conditions. Some applications of this theory are presented only with the connection between the auxiliary problem and a "dual" of the given problem.

DD FORM 1 NOV 65 1473

Unclassified
Security Classification

Unclassified

Security Classification

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
concave differentiable function constrained optimization problems Lagrange multiplier mathematical programming nonlinear programming problems						

Unclassified

Security Classification